GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 14, 2001, 14:37:58 ; Search time 40.49 Seconds
(without alignments)
1421.405 Million cell updates/sec

Title: PCT-US01-18568A-2
Perfect score: 2342
Sequence: 1 MDPDSDQPLNSLDVKPLRKP.....VYTKVSAYLNWIYNVWKAEL 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_16:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2337	99.8	437	4	Q9NRS4	Q9nrs4 homo sapien
2	2198	93.9	423	4	Q9NZA5	Q9nza5 homo sapien
3	686.5	29.3	767	13	Q9DGR2	Q9dgr2 xenopus lae
4	628.5	26.8	445	11	Q9ER03	Q9er03 mus musculu
5	574	24.5	457	4	Q9H3S3	Q9h3s3 homo sapien
6	568.5	24.3	311	11	Q9ER04	Q9er04 mus musculu
7	568	24.3	418	4	O60235	O60235 homo sapien
8	562	24.0	273	11	Q9ER02	Q9er02 mus musculu
9	527	22.5	279	11	Q9QZ74	Q9qz74 rattus norv
10	504.5	21.5	329	6	Q9GL10	Q9gl10 ovis aries
11	504.5	21.5	1524	13	Q91674	Q91674 xenopus lae
12	503.5	21.5	305	11	Q9JHJ7	Q9jhj7 mus musculu
13	498.5	21.3	377	6	P79343	P79343 bos taurus
14	496	21.2	643	6	O97506	O97506 sus scrofa
15	493.5	21.1	421	11	Q60491	Q60491 cavia porce
16	488.5	20.9	1113	11	Q9Z319	Q9z319 mus musculu
17	488	20.8	812	11	Q9R0W3	Q9r0w3 rattus norv
18	486	20.8	389	13	Q9PVX7	Q9pvx7 xenopus lae
19	484	20.7	1042	4	Q9Y5Q5	Q9y5q5 homo sapien
20	483	20.6	421	4	Q9UBB2	Q9ubb2 homo sapien
21	479.5	20.5	845	13	Q9DGR1	Q9dgr1 xenopus lae
22	479	20.5	321	4	Q9NRQ8	Q9nrq8 homo sapien
23	478.5	20.4	415	6	Q29015	Q29015 sus scrofa
24	478	20.4	314	4	Q9Y6M0	Q9y6m0 homo sapien
25	477	20.4	1379	5	Q9V4N6	Q9v4n6 drosophila
26	476.5	20.3	311	11	Q9QUL7	Q9qul7 mus musculu
27	476	20.3	581	5	Q9XZM7	Q9xzm7 strongyloce
28	475	20.3	422	4	Q9UL52	Q9ul52 homo sapien
29	474	20.2	855	11	Q9JJI7	Q9jji7 rattus norv
30	473.5	20.2	260	13	Q9W7P9	Q9w7p9 paralichthy
31	473.5	20.2	310	11	Q9QYZ9	Q9qyz9 mus musculu
32	472.5	20.2	317	4	Q9GZN4	Q9gzn4 homo sapien
33	472	20.2	312	4	Q9NS34	Q9ns34 homo sapien
34	469	20.0	321	4	Q9NRR2	Q9nrr2 homo sapien
35	467	19.9	855	4	Q9H3S0	Q9h3s0 homo sapien
36	464	19.8	266	13	Q9W7Q0	Q9w7q0 paralichthy
37	463	19.8	855	4	Q9HCA3	Q9hca3 homo sapien
38	459.5	19.6	342	11	Q9ES87	Q9es87 rattus norv
39	459.5	19.6	342	11	Q9ER01	Q9er01 rattus norv
40	459.5	19.6	1322	5	Q9NAT0	Q9nat0 anopheles g
41	457.5	19.5	571	4	Q9Y495	Q9y495 homo sapien
42	456.5	19.5	1322	5	Q9NJS5	Q9njs5 anopheles q
43	456	19.5	868	5	Q9Y1V1	Q9y1v3 polyandroca
44	454.5	19.4	256	5	Q9XYY1	Q9xyy1 rhyzopertha
45	452.5	19.3	339	11	Q9ESD1	Q9esd1 mus musculu

ALIGNMENTS

RESULT 1
Q9NRS4
ID Q9NRS4 PRELIMINARY; PRT; 437 AA.
AC Q9NRS4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE TRANSMEMBRANE SERINE PROTEASE 3.
GN TMPRSS3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREATIC CARCINOMA;
RX MEDLINE=20283276; PubMed=10825129;
RA Wallrapp C., Hahnel S., Muller-Pillasch F., Burghardt B., Iwamura T.,
RA Ruthenburger M., Lerch M.M., Adler G., Gress T.M.;
RT "A novel transmembrane serine protease (TMPRSS3) overexpressed in
RT pancreatic cancer.";
RL Cancer Res. 60:2602-2606(2000).
CC -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR EMBL; AF179224; AAF74526.1; -.
DR InterPro; IPR001254; -.
DR InterPro; IPR001314; -.
DR InterPro; IPR002172; -.
DR Pfam; PF00057; ldl_recept_a; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR SMART; SM00192; LDLa; 1.
KW Protease.
SQ SEQUENCE 437 AA; 48204 MW; 351B2FD4AB657B12 CRC64;

Query Match 99.8%; Score 2337; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 7.2e-211;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALLSLASIIIVVVLIKVILDKYYFL 61

```
Db      4 DPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALLSLASIIIVVVLIKVILDKYYFL 63

Qy     62 CGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNW 121

Db     64 CGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNW 123

Qy    122 FSACFDNFTEALAETACRQMGYSSKPTFRAVEIGPDQDLDVVEITENSQELRMRNSSGPC 181

Db    124 FSACFDNFTEALAETACRQMGYSSKPTFRAVEIGPDQDLDVVEITENSQELRMRNSSGPC 183

Qy    182 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPWQVSIQYDKQHVCGGSILDPHWVLTA 241

Db    184 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPWQVSIQYDKQHVCGGSILDPHWVLTA 243

Qy    242 AHCFRKHTDVFNWKVRAGSDKLGSFPSLAVAKIIIIEFNPMYPKDNDIALMKLQFPLTFS 301

Db    244 AHCFRKHTDVFNWKVRAGSDKLGSFPSLAVAKIIIIEFNPMYPKDNDIALMKLQFPLTFS 303

Qy    302 GTVRPICLPFFDEELTPATPLWIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQ 361

Db    304 GTVRPICLPFFDEELTPATPLWIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQ 363

Qy    362 GEVTEKMMCAGIPEGGVDTCQGDSGGPLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVS 421

Db    364 GEVTEKMMCAGIPEGGVDTCQGDSGGPLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVS 423

Qy    422 AYLNWIYNVWKAEL 435

Db    424 AYLNWIYNVWKAEL 437
```

```
RESULT   2
Q9NZA5
ID   Q9NZA5      PRELIMINARY;      PRT;   423 AA.
AC   Q9NZA5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   TYPE II MEMBRANE SERINE PROTEASE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Smeekens S.S., Lorimar D.D., Wang E., Hou J., Linnevers C.;
RT   "MT-SP2, a novel type II membrane serine protease expressed in
RT   trachea, colon, and small intestine: identification, cloning, and
RT   chromosomal localization.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR   EMBL; AF216312; AAF31436.1; -.
DR   InterPro; IPR001254; -.
DR   InterPro; IPR001314; -.
DR   InterPro; IPR002172; -.
DR   Pfam; PF00057; ldl_recept_a; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
DR   SMART; SM00192; LDLa; 1.
KW   Protease.
SQ   SEQUENCE   423 AA;  46397 MW;  90792AF0F8AFEA3Q CRC64;

 Query Match           93.9%;  Score 2198;  DB 4;  Length 423;
 Best Local Similarity 97.4%;  Pred. No. 7.6e-198;
 Matches  ...  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 MKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIGWGFTKQNGGKMSDILLQASVQVIDS 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIGWGFTKQNGGKMSDILLQASVQVIDS 60

Qy     61 CGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNW 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNW 120

Qy    121 FSACFDNFTEALAETACRQMGYSSKPTFRAVEIGPDQDLDVVEITENSQELRMRNSSGPC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 FSACFDNFTEALAETACRQMGYSSKPTFRAVEIGPDQDLDVVEITENSQELRMRNSSGPC 180

Qy    181 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPWQVSIQYDKQHVCGGSILDPHWVLTA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPWQVSIQYDKQHVCGGSILDPHWVLTA 240

Qy    241 AHCFRKHTDVFNWKVRAGSDKLGSFPSLAVAKIIIIEFNPMYPKDNDIALMKLQFPLTFS 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AHCFRKHTDVFNWKVRAGSDKLGSFPSLAVAKIIIIEFNPMYPKDNDIALMKLQFPLTFS 300

Qy    301 GTVRPICLPFFDEELTPATPLWIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQ 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GTVRPICLPFFDEELTPATPLWIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQ 360

Qy    361 GEVTEKMMCAGIPEGGVDTCQGDSGGPLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GEVTEKMMCAGIPEGGVDTCQGDSGGPLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVS 420

Qy    421 AYLNWIYNVWKAEL 434
          ||||||||||||||
Db    421 AYLNWIYNVWKAEL 434
```

```
Qy     76 DGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDNFTEALAE 135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 DGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDNFTEALAE 123

Qy    136 TACRQMGYSSKPTFRAVEIGPDQDLDVVEITENSQELRMRNSSGPCLSGSLVSLHCLACG 195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 TACRQMGYSSKPTFRAVEIGPDQDLDVVEITENSQELRMRNSSGPCLSGSLVSLHCLACG 183

Qy    196 KSLKTPRVVGGEEASVDSWPWQVSIQYDKQHVCGGSILDPHWVLTAAHCFRKHTDVFNWK 255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 KSLKTPRVVGGEEASVDSWPWQVSIQYDKQHVCGGSILDPHWVLTAAHCFRKHTDVFNWK 243

Qy    256 VRAGSDKLGSFPSLAVAKIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEE 315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 VRAGSDKLGSFPSLAVAKIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEE 303

Qy    316 LTPATPLWIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPE 375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 LTPATPLWIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPE 363

Qy    376 GGVDTCQGDSGGPLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL 435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 GGVDTCQGDSGGPLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL 423
```

```
RESULT   3
Q9DGR2
ID   Q9DGR2      PRELIMINARY;      PRT;   767 AA.
AC   Q9DGR2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   EMBRYONIC SERINE PROTEASE-2.
GN   XESP-2.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20363741; PubMed=10903452;
RA   Yamada K., Takabatake T., Takeshima K.;
RT   "Isolation and characterization of three novel serine protease genes
RT   from Xenopus laevis.";
RL   Gene 252:209-216(2000).
DR   EMBL; AB038497; BAB08217.1; -.
KW   Protease.
SQ   SEQUENCE   767 AA;  86001 MW;  E0566A38796DE96E CRC64;

 Query Match           29.3%;  Score 686.5;  DB 13;  Length 767;
 Best Local Similarity 38.3%;  Pred. No. 1.1e-55;
 Matches  145;  Conservative   58;  Mismatches  153;  Indels   23;  Gaps    7;

Qy     62 CGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNW 121
          ||  :  :   ( |||  ||| |||| ||  :|                 |||  ::   |
Db    395 CGSSVSCVLSSQWCDGVSDCPYGEDEMSCVSLYPAD----------FQLQVYSTSVSAW 443

Qy    122 FSACFDNFTEALAETACRQMGYSSKPTFRAVEI----GPDQDLDVVEITENSQELRMRNS 177
            | | : :     ||:  ||:     |   :     |:    :      |:
Db    444 LPVCSDYWNDDFGRFACQDFGYNGSSYNRYDTLMSPYAPNGYFKLYSGYWRSKFYTSVQY 503

Qy    178 SGPCLSGSLVSLHCLACGKSLKT--PRVVGGEEASVDSWPWQVSIQYDKQHVCGGSILDP 235
          |  | ||::|||||::|| |  :   |:|||  |:: :|||||::||    :|||||: |
Db    504 SSYCYSGNVVSLHCISCGVSNNSLVSRIVGGTFANLGNWPWQVNLQYITGVLCGGSIISP 563

Qy    236 HWVLTAAHC-FRKHTDVFNWKVRAGSDKLGSFPSLA---VAKIIIIEFNPMYPKDNDIAL 291
           |::|||||  :  ::    ):| ||:    |: : :   | :||:      |  ||||||
Db    564 KWIVTAAHCVYGSYSSASGWRVFAGTLTKPSYYNASAYFVERIIVHPGYKSYTYDNDIAL 623

Qy    292 MKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIGWGFTKQNGGKMSDILLQASVQVIDS 351
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 14, 2001, 14:39:05 ; Search time 34.69 Seconds
(without alignments)
510.297 Million cell updates/sec

Title: PCT-US01-18568A-9
Perfect score: 1600
Sequence: 1 MDSKGSSQKSRLLLLLVVSN.....LNWIYNVWKAELSRHHHHHH 292

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601:*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1266	79.1	432	21	AAY99417	Human PRO1570 (UNQ
2	1266	79.1	432	22	AAB87581	Human PRO1570. Ho
3	1266	79.1	432	22	AAB66166	Protein of the inv
4	1266	79.1	435	20	AAY06437	Human protease HUP
5	1266	79.1	435	22	AAY72558	Human seripancrin
6	1264	79.0	233	18	AAW22986	Human serine prote
7	1253	78.3	492	22	AAY72559	Human seripancrin
8	1230	76.9	240	21	AAB23166	Human colorectal c
9	754.5	47.2	315	22	AAB73946	Fusion protein of
10	746	46.6	328	21	AAB36480	Fusion gene with h
11	746	46.6	328	22	AAB67540	Amino acid sequenc
12	740.5	46.3	327	22	AAY72891	PFEK-C-E-HIS fusio
13	672.5	42.0	289	21	AAB36483	Fusion gene with h
14	672.5	42.0	289	22	AAB67543	Amino acid sequenc
15	645	40.3	288	21	AAB36482	Fusion gene with h
16	645	40.3	288	22	AAB67542	Amino acid sequenc
17	601.5	37.6	306	22	AAB67544	Amino acid sequenc
18	593	37.1	319	21	AAB36481	Fusion gene with h
19	593	37.1	319	22	AAB67541	Amino acid sequenc
20	574	35.9	284	22	AAB67545	Amino acid sequenc
21	528	33.0	296	21	AAY72108	Human serine prote
22	528	33.0	372	21	AAY72092	Human serine prote
23	526.5	32.9	418	17	AAR89435	Trypsin-like enzym
24	526.5	32.9	418	20	AAY29501	Human lung tumour
25	526.5	32.9	418	20	AAY29502	Human lung tumour
26	526.5	32.9	418	21	AAB44437	Human lung tumour-
27	526.5	32.9	418	21	AAB44438	Human lung tumour-
28	524.5	32.8	418	20	AAY29498	Human lung tumour
29	524.5	32.8	418	21	AAB44428	Human lung tumour-
30	522.5	32.7	232	17	AAR89430	Trypsin-like enzym
31	517	32.3	273	21	AAB11696	Mouse serine prote
32	517	32.3	311	21	AAB11697	Mouse serine prote
33	517	32.3	445	21	AAB11698	Mouse serine prote
34	515	32.2	238	21	AAB11695	Mouse serine prote
35	509.5	31.8	248	21	AAB43572	Human cancer assoc
36	509.5	31.8	327	21	AAY72093	Human serine prote
37	508	31.8	492	21	AAB36901	Human TMPRSS2 prot
38	508	31.8	492	21	AAY57280	Ovr115 homolog pro
39	507	31.7	452	20	AAY41694	Human PRO382 prote
40	507	31.7	453	21	AAB44250	Human PRO382 (UNQ3
41	507	31.7	454	21	AAB32246	Tumour associated
42	506	31.6	492	21	AAY92050	HrPCa6/7 polypepti
43	506	31.6	492	21	AAY77726	Human tumour suppr
44	506	31.6	492	21	AAY44406	Human 20P1F12-GTC2
45	497	31.1	283	21	AAY81492	Human prostate-ass

ALIGNMENTS

RESULT 1
AAY99417
ID AAY99417 standard; Protein; 432 AA.
XX
AC AAY99417;
XX
DT 08-AUG-2000 (first entry)
XX
DE Human PRO1570 (UNQ776) amino acid sequence SEQ ID NO:275.
XX
KW Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening.
XX
OS Homo sapiens.
XX
PN WO200012708-A2.
XX
PD 09-MAR-2000.
XX
PF 01-SEP-1999; 99WO-US20111.
XX
PR 01-SEP-1998; 98US-0098716.
PR 01-SEP-1998; 98US-0098749.
PR 01-SEP-1998; 98US-0098750.
PR 02-SEP-1998; 98US-0098803.
PR 02-SEP-1998; 98US-0098821.
PR 02-SEP-1998; 98US-0098843.
PR 09-SEP-1998; 98US-0099536.
PR 09-SEP-1998; 98US-0099596.
PR 09-SEP-1998; 98US-0099598.
PR 09-SEP-1998; 98US-0099602.
PR 09-SEP-1998; 98US-0099642.
PR 10-SEP-1998; 98US-0099741.
PR 10-SEP-1998; 98US-0099754.

PR 10-SEP-1998; 98US-0099763.
PR 10-SEP-1998; 98US-0099792.
PR 10-SEP-1998; 98US-0099808.
PR 10-SEP-1998; 98US-0099812.
PR 10-SEP-1998; 98US-0099815.
PR 10-SEP-1998; 98US-0099816.
PR 15-SEP-1998; 98US-0100385.
PR 15-SEP-1998; 98US-0100388.
PR 15-SEP-1998; 98US-0100390.
PR 16-SEP-1998; 98US-0100584.
PR 16-SEP-1998; 98US-0100627.
PR 16-SEP-1998; 98US-0100661.
PR 16-SEP-1998; 98US-0100662.
PR 16-SEP-1998; 98US-0100664.
PR 17-SEP-1998; 98US-0100683.
PR 17-SEP-1998; 98US-0100684.
PR 17-SEP-1998; 98US-0100710.
PR 17-SEP-1998; 98US-0100711.
PR 17-SEP-1998; 98US-0100919.
PR 17-SEP-1998; 98US-0100930.
PR 18-SEP-1998; 98US-0100848.
PR 18-SEP-1998; 98US-0100849.
PR 18-SEP-1998; 98US-0101014.
PR 18-SEP-1998; 98US-0101068.
PR 18-SEP-1998; 98US-0101071.
PR 22-SEP-1998; 98US-0101279.
PR 23-SEP-1998; 98US-0101471.
PR 23-SEP-1998; 98US-0101472.
PR 23-SEP-1998; 98US-0101474.
PR 23-SEP-1998; 98US-0101475.
PR 23-SEP-1998; 98US-0101476.
PR 23-SEP-1998; 98US-0101477.
PR 23-SEP-1998; 98US-0101479.
PR 24-SEP-1998; 98US-0101738.
PR 24-SEP-1998; 98US-0101741.
PR 24-SEP-1998; 98US-0101743.
PR 24-SEP-1998; 98US-0101915.
PR 24-SEP-1998; 98US-0101916.
PR 29-SEP-1998; 98US-0102207.
PR 29-SEP-1998; 98US-0102240.
PR 29-SEP-1998; 98US-0102307.
PR 29-SEP-1998; 98US-0102330.
PR 29-SEP-1998; 98US-0102331.
PR 30-SEP-1998; 98US-0102484.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102570.
PR 30-SEP-1998; 98US-0102571.
PR 01-OCT-1998; 98US-0102684.
PR 01-OCT-1998; 98US-0102687.
PR 02-OCT-1998; 98US-0102965.
PR 06-OCT-1998; 98US-0103258.
PR 06-OCT-1998; 98US-0103449.
PR 07-OCT-1998; 98US-0103314.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103395.
PR 07-OCT-1998; 98US-0103396.
PR 07-OCT-1998; 98US-0103401.
PR 08-OCT-1998; 98US-0103633.
PR 08-OCT-1998; 98US-0103678.
PR 08-OCT-1998; 98US-0103679.
PR 08-OCT-1998; 98US-0103711.
PR 14-OCT-1998; 98US-0104257.
PR 20-OCT-1998; 98US-0104987.
PR 20-OCT-1998; 98US-0105000.
PR 20-OCT-1998; 98US-0105002.
PR 21-OCT-1998; 98US-0105104.
PR 22-OCT-1998; 98US-0105169.
PR 22-OCT-1998; 98US-0105266.
PR 26-OCT-1998; 98US-0105693.
PR 26-OCT-1998; 98US-0105694.
PR 27-OCT-1998; 98US-0105807.
PR 27-OCT-1998; 98US-0105881.
PR 27-OCT-1998; 98US-0105882.
PR 27-OCT-1998; 98US-0106062.
PR 28-OCT-1998; 98US-0106023.
PR 28-OCT-1998; 98US-0106029.
PR 28-OCT-1998; 98US-0106030.
PR 28-OCT-1998; 98US-0106032.
PR 28-OCT-1998; 98US-0106033.
PR 28-OCT-1998; 98US-0106178.
PR 29-OCT-1998; 98US-0106248.
PR 29-OCT-1998; 98US-0106384.
PR 29-OCT-1998; 98US-0108500.
PR 30-OCT-1998; 98US-0106464.
PR 03-NOV-1998; 98US-0106856.
PR 03-NOV-1998; 98US-0106902.
PR 03-NOV-1998; 98US-0106905.
PR 03-NOV-1998; 98US-0106919.
PR 03-NOV-1998; 98US-0106932.
PR 03-NOV-1998; 98US-0106934.
PR 10-NOV-1998; 98US-0107783.
PR 17-NOV-1998; 98US-0108775.
PR 17-NOV-1998; 98US-0108779.
PR 17-NOV-1998; 98US-0108787.
PR 17-NOV-1998; 98US-0108788.
PR 17-NOV-1998; 98US-0108801.
PR 17-NOV-1998; 98US-0108802.
PR 17-NOV-1998; 98US-0108806.
PR 17-NOV-1998; 98US-0108807.
PR 17-NOV-1998; 98US-0108867.
PR 17-NOV-1998; 98US-0108925.
PR 18-NOV-1998; 98US-0108848.
PR 18-NOV-1998; 98US-0108849.
PR 18-NOV-1998; 98US-0108850.
PR 18-NOV-1998; 98US-0108851.
PR 18-NOV-1998; 98US-0108852.
PR 18-NOV-1998; 98US-0108858.
PR 18-NOV-1998; 98US-0108904.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
XX
DR WPI; 2000-237871/20.
DR N-PSDB; AAA37099.
XX
PT New mammalian DNA sequences encoding transmembrane, receptor or
PT secreted PRO polypeptides, useful for screening of potential peptide or
PT small molecule inhibitors of the relevant receptor/ligand interactions
XX
PS Claim 12; Fig 156; 773pp; English.
XX
CC AAA37022 to AAA37144 encode the new isolated human transmembrane,
CC receptor or secreted PRO polypeptides given in AAY99340 to AAY99462. The
CC transmembrane and receptor PRO proteins can be used for screening of
CC potential peptide or small molecule inhibitors of the relevant
CC receptor/ligand interactions. The polypeptides and nucleotide sequences
CC encoding then have various industrial applications, including uses as
CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent
CC PCR primers and hybridisation probes used in the isolation of the PRO
CC polypeptides from the present invention.
XX
SQ Sequence 432 AA;

Query Match 79.1%; Score 1266; DB 21; Length 432;
Best Local Similarity 97.4%; Pred. No. 1.9e-112;
Matches 228; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 51 KIVGGYALDVDSWPWQVSIQYDKQHVCGGSILDPHWVLTAAHCFRKHTDVFNWKVRAGSD 110
::||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 199 rvvggeeasvdswpwqvsiqydkqhvcggsildphwvltaahcfrkhtdvfnwkvragsd 258

Qy 111 KLGSFPSLAVAKIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATP 170
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db    259 klgsfpslavakiiiiefnpmypkdndialmklqfpltfsgtvrpiclpffdeeltpatp 318

Qy    171 LWIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTC 230
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    319 lwiigwgftkqnggkmsdillqusvqvidstrcnaddayqgevtekmmcagipeggvdtc 378

Qy    231 QGDSGGPLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL 284
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    379 qgdsggplmyqsdqwhvvgivswgygcggpstpgvytkvsaylnwiynvwkael 432
```

RESULT 2
AAB87581
ID AAB87581 standard; Protein; 432 AA.
XX
AC AAB87581;
XX
DT 15-MAY-2001 (first entry)
XX
DE Human PRO1570.
XX
KW Human; PRO protein; mapping.
XX
OS Homo sapiens.
XX
PN WO200116318-A2.
XX
PD 08-MAR-2001.
XX
PF 24-AUG-2000; 2000WO-US23328.
XX
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 07-DEC-1999; 99US-0169495.
PR 09-DEC-1999; 99US-0170262.
PR 11-JAN-2000; 2000US-0175481.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 01-MAR-2000; 2000WO-US05601.
PR 03-MAR-2000; 2000US-0187202.
PR 25-APR-2000; 2000US-0199397.
PR 22-MAY-2000; 2000WO-US14042.
PR 05-JUN-2000; 2000US-0209832.
XX
PA (GETH) GENENTECH INC.
XX
PI Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;
XX
DR WPI; 2001-183260/18.
DR N-PSDB; AAF92113.
XX
PT Eighty four nucleic acids encoding PRO polypeptides, useful in
PT molecular biology, including use as hybridization probes, and in
PT chromosom, and gene mapping. -
XX
PS Claim 12; Fig 112; 278pp; English.
XX
CC The present sequence is a human PRO polypeptide (secreted and
CC transmembrane). The PRO protein, and PRO agonists, PRO antagonists or
CC anti-PRO antibodies are useful for preparation of a medicament useful in
CC the treatment of a condition which is responsive to the PRO protein,
CC agonists, antagonists or anti-PRO antibodies. The PRO protein may also be
CC employed as molecular weight markers for protein electrophoresis. The PRO
CC coding sequence has applications in molecular biology, including use as
CC hybridisation probes, and in chromosome and gene mapping.
XX
SQ Sequence 432 AA;

Query Match 79.1%; Score 1266; DB 22; Length 432;
Best Local Similarity 97.4%; Pred. No. 1.9e-112;

Matches 228; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy     51 KIVGGYALDVDSWPWQVSIQYDKQHVCGGSILDPHWVLTAAHCFRKHTDVFNWKVRAGSD 110
          ::|||    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199 rvvggeeasvdswpwqvsiqydkqhvcggsildphwvltaahcfrkhtdvfnwkvragsd 258

Qy    111 KLGSFPSLAVAKIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATP 170
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259 klgsfpslavakiiiiefnpmypkdndialmklqfpltfsgtvrpiclpffdeeltpatp 318

Qy    171 LWIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTC 230
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    319 lwiigwgftkqnggkmsdillqasvqvidstrcnaddayqgevtekmmcagipeggvdtc 378

Qy    231 QGDSGGPLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL 284
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    379 qgdsggplmyqsdqwhvvgivswgygcggpstpgvytkvsaylnwiynvwkael 432
```

RESULT 3
AAB66166
ID AAB66166 standard; protein; 432 AA.
XX
AC AAB66166;
XX
DT 02-APR-2001 (first entry)
XX
DE Protein of the invention #78.
XX
KW Secreted; transmembrane; gene therapy.
XX
OS Unidentified.
XX
PN WO200078961-A1.
XX
PD 28-DEC-2000.
XX
PF 18-FEB-2000; 2000WO-US04342.
XX
PR 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 01-SEP-1999; 99WO-US20111.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28551.
PR 16-DEC-1999; 99WO-US30095.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00376.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
PI Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D;
PI Watanabe CK, Williams PM, Wood WI;
XX
DR WPI; 2001-071395/08.
XX
PT Secreted and transmembrane proteins and nucleic acids designated PRO,
PT useful as hybridization probes, in chromosome and gene mapping and gene
PT therapy -
XX
PS Claim 1; Fig 156; 787pp; English.
XX
CC The present invention relates to secreted and transmembrane proteins.
CC These proteins and the DNA encoding them may be used as hybridization
CC probes, in chromosome and gene mapping and in the generation of
CC anti-sense RNA and DNA. They may also be used used to generate either
CC transgenic animals or knockout animals which are in turn useful for
CC development and screening of therapeutically useful reagents.
CC The nucleic acids may also be used in gene therapy.
XX